Supplementary Information for:
A Systematic Study of DNA Conformation in Slitlike Confinement

Liang Dai[a], Jeremy J. Jones[b], Johan R.C. van der Maarel[c], and Patrick S. Doyle[a,b*]

[a] *BioSystems and Micromechanics (BioSyM) IRG, Singapore-MIT Alliance for Research and Technology (SMART) Centre, 3 Science Drive 2, Republic of Singapore 117543*
[b] *Department of Chemical Engineering, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139*
[c] *Department of Physics, National University of Singapore, 2 Science Drive 3, Republic of Singapore 117542*
[*] *email: pdoyle@mit.edu*

Excluded volume interaction for a DNA chain confined to a plane

The excluded volume interaction for a DNA chain confined to a plane is determined by the excluded area of a DNA segment instead of the excluded volume of a segment, because the system is two-dimensional. We consider the excluded area between two straight segments (Figure S1). The segment has a length of l and a width of w. The area inside the red polygon (Figure S1) is the excluded area of the light grey segment, when the skewed angle is fixed as θ. Assuming the orientations of DNA segments are random, the average excluded area for two straight DNA segments is calculated as:

$$A_{ev} = \frac{2}{\pi}\int_0^{\pi/2}\left[l^2\sin\theta + w^2\sin\theta + 4lw\cos^2(\theta/2)\right]d\theta$$
$$= \frac{2}{\pi}\left[l^2 + (2+\pi)lw + w^2\right] \qquad \text{(S1)}$$
$$= \frac{2}{\pi}\left[l + (1+\frac{\pi}{2})w\right]^2 - (\frac{\pi}{2}+2)w^2$$

Usually, DNA is considered stiff over a single Kuhn length, which is two times the persistence length L_p. So we set $l = 2L_p$ in Eqn. (S1). In the case of $w \ll l$, the second term in Eqn. (S1) can be ignored. After ignoring the prefactor, the excluded area follows:

$$A_{ev} \sim \left[L_p + (\frac{1}{2}+\frac{\pi}{4})w\right]^2 \approx \left(L_p + 1.3w\right)^2 \qquad \text{(S2)}$$

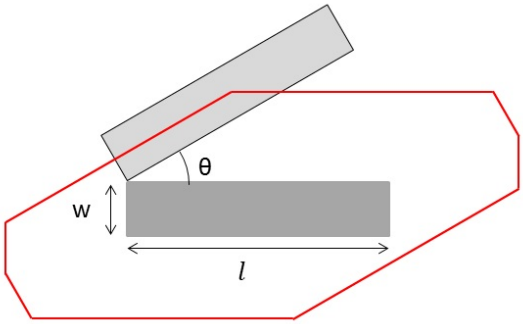

Figure S1. Illustration of the excluded area of the light grey segment by the dark grey segment with a fixed skewed angle θ. The red polygon is the excluded area. The segment has a length of l and a width of w.

Benchmarking the simulation using large bond length l_B

Setting the large value of bond length l_B in the simulation can reduce the computational time for a given contour length L, because the computational time is mainly determined by the number of beads $N_{bead} = L/l_B + 1$. However, using large l_B in simulation will produce some problems. First, the value of l_B cannot be larger than the chain width w. Otherwise, there will be a gap between two adjacent beads and the chain in the simulation is not solid-filled. The gap

between two adjacent beads can cause the chain overlapping, which is artificial and should be avoided. Figure S2(a) illustrates a solid-filled chain. Second, a simulation using large l_B may not precisely capture the DNA bending. This concept of insufficient fine graining of the DNA chain is illustrated in Figure S2(b). In addition, the DNA conformation in a narrow slit cannot be precisely modeled when using a bond length larger than the slit height *H*. So, two conditions, $l_B \leq w$ and $l_B < H$, should be satisfied in the parameter setting.

Using large l_B affects the contour length of DNA in the simulation. It is because the series of bonds in simulation is not exactly the wormlike chain with smooth bending as in the real case (Figure S2(b)). Each straight segment of the red line with a length of l_B corresponds to a bending segment of the gray line with a contour length larger than l_B, which is referred as l_B'. Accordingly, the contour length of the wormlike chain is $(N_{bead}-1)l_B'$, rather than $(N_{bead}-1)l_B$. The relationship between l_B' and l_B follows[1]:

$$(l_B')^2 = 2L_p^2\left[\frac{l_B}{L_p} - 1 + \exp(-l_B / L_p)\right]. \tag{S3}$$

Applying Eqn. (S3), we obtain that when l_B is 5, 10, 20 or 40 nm, l_B' is 5.08, 10.34 21.43 or 46.12 nm.

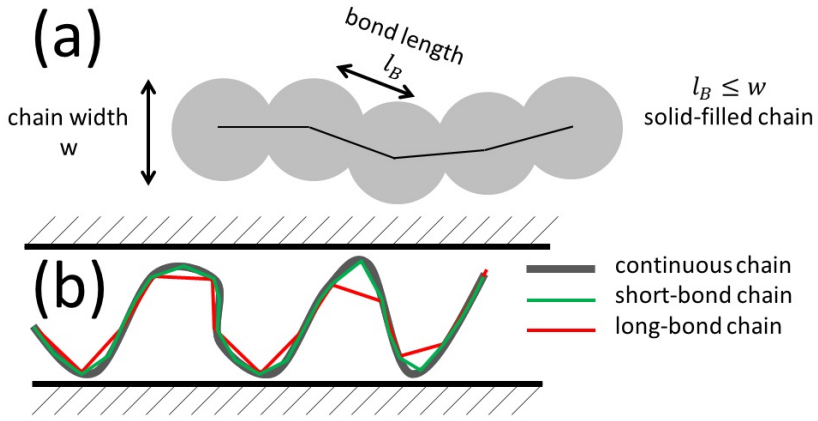

Figure S2. (a) Illustration of bead-bond model. There is no gap between two beads if $l_B \leq w$. Usually, we set $l_B = w$ in simulations. (b) A continuous wormlike chain is discretized to a series of bonds. A short-bond chain with $l_B \ll L_p$ can precisely model the smooth bending. A long-bond chain does not precisely model the bending, especially on the length scale less than l_B.

Furthermore, using large l_B also affects the setting of the bending rigidity. In the simulation, we calculate the bending energy by:

$$\frac{E_{i,i+1}^{bend}(\theta_{i,i+1})}{k_bT} = \frac{1}{2}\frac{\kappa}{l_B}\theta_{i,i+1}^2 \approx \frac{1}{2}\frac{L_p}{l_B}\theta_{i,i+1}^2 \tag{S4}$$

,

where κ is the bending rigidity, and $\theta_{i,i+1}$ is the bending angle between the bond *i* and *i+1*. Usually, if $l_B \ll L_p$, κ in Eqn. (S4) is equal to L_p. However, when l_B is comparable to L_p, the term L_p/l_B in Eqn. (S4) should be replaced by $L_p/\langle l_B'\rangle$ to precisely capture the persistence length of the wormlike chain. Considering that $l_B < \langle l_B'\rangle$ and $L_p/l_B > L_p/\langle l_B'\rangle$, we can infer that using Eqn. (S4) with $\kappa = L_p$ will overestimate the persistence length. In the current study, we always set the bending rigidity $\kappa = L_p$ for simplification. To quantify the overestimation of the persistence length due to using large l_B, we perform several benchmark simulations of DNA in bulk without excluded volume interaction. The precise value of L_p of DNA in the simulation can be obtained by fitting Benoit-Doty equation[2].

$$\left\langle R_g^2\right\rangle = \frac{L_p}{3L^2}\left[L^3 - 3L^2L_p + 6LL_p^2 - 6L_p^3(1-\exp(-\frac{L}{L_p}))\right] \tag{S5}$$

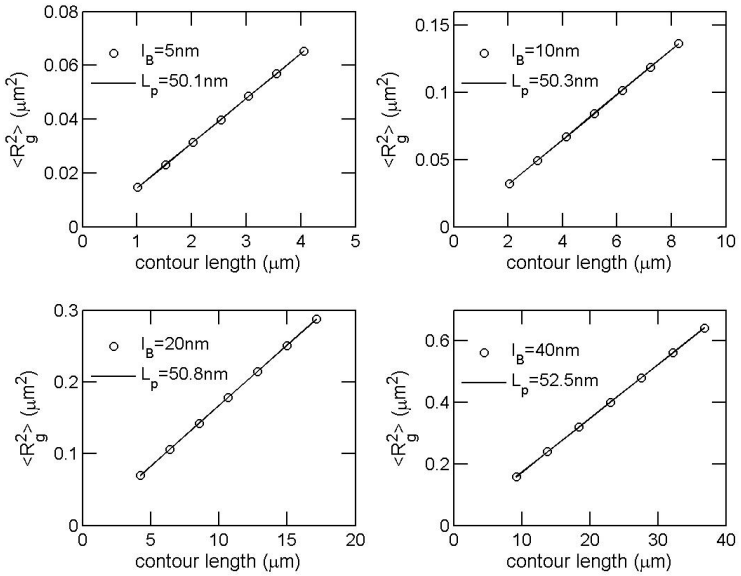

Figure S3. Mean squares of the radius of gyration as a function of the contour length obtained from the simulations using different bond length l_B. The simulations are in bulk without excluded volume interaction. Open circles are simulation results using the same bending rigidity κ = 50 nm in Eqn. (S4) but using different l_B. The solid line is calculated from Eqn. (S5) using a fitted L_p. The fitted values and uncertainties are 50.1 ± 0.1, 50.3 ± 0.1, 50.8 ± 0.1, 52.5 ± 0.1 nm, respectively.

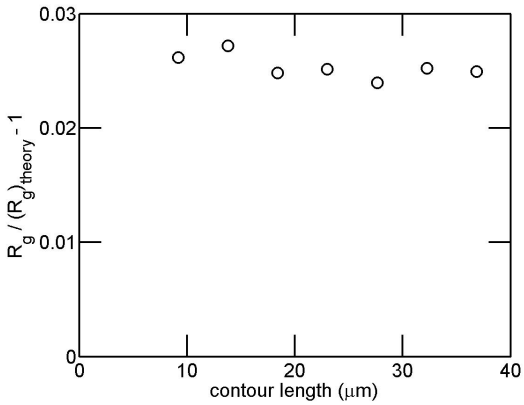

Figure S4. Deviation of R_g in simulation from the theoretical value as a function of the contour length. The simulations are performed in bulk without excluded volume interaction. The bond length l_B is 40 nm.

Figure S3 shows $\langle R_g^2 \rangle$ as a function as L. The simulation results are fitted by Eqn. (S5) with a fitted L_p. As expected, the fitted value of L_p is slightly larger than 50 nm and the deviation become more significant for larger l_B. The largest l_B in the current study is 40 nm, which causes 5% overestimation of the persistence length compared to the value assumed in the simulation. In the current study, we are interested in the value of R_g. Figure S4 shows the relative overestimation of R_g when we set l_B in the simulations. It indicates the overestimation is always less than 3%.

Above simulations are performed for DNA in bulk without excluded volume interaction. To examine the effect of using different l_B in the simulation of a real chain in a slit, we performed two sets of simulation using same parameters except different bond lengths, as shown in Figure S5. The results from two sets of simulation agree with each other except the region of $H < l_B$. Such result demonstrates that using large l_B does not affect the simulation result in a wide slit.

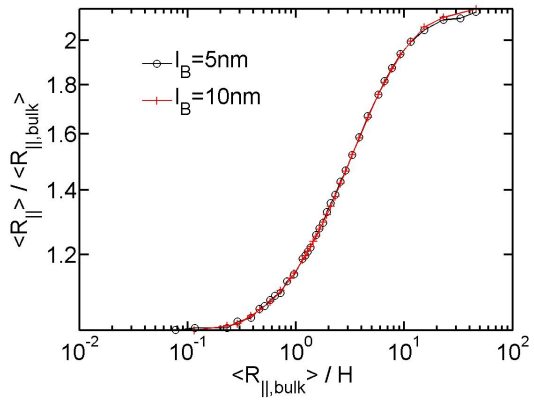

Figure S5. Comparison of two simulation results using the bond length of 5 nm and 10 nm, respectively. $\langle R_{||} \rangle$ and $\langle R_{||,bulk} \rangle$ are the average in-plane radius of gyration in confinement and bulk, respectively. *H* is the slit height. Both simulations share the contour length of 4 μm, the chain width of 10 nm and the bending rigidity κ = 50 nm in Eqn. (S4).

References

1. J. R. C. van der Maarel, *Introduction To Biopolymer Physics*, World Scientific, Singapore, 2007.
2. H. Benoit and P. Doty, *J. Phys. Chem.*, 1953, **57**, 958-963.